Jointly published by Akadémiai Kiadó, Budapest and Kluwer Academic Publishers, Dordrecht

Scientometrics, Vol. 56, No. 2 (2003) 259–282

Correcting glasses help fair comparisons in international science landscape: Country indicators as a function of ISI database delineation

MICHEL ZITT,[1,2] SUZY RAMANANA-RAHARY,[1] ELISE BASSECOULARD[2]

[1]*Observatoire des Sciences et des Techniques (OST), Paris (France)*
[2]*UMR EDRA, INRA-Université de Nantes, Nantes (France)*

The increasing use of bibliometric indicators in science policy calls for a reassessment of their robustness and limits. The perimeter of journal inclusion within ISI databases will determine variations in the classic bibliometric indicators used for international comparison, such as world shares of publications or relative impacts. We show in this article that when this perimeter is adjusted using a natural criterion for inclusion of journals, the journal impact, the variation of the most common country indicators (publication and citation shares; relative impacts) with the perimeter chosen depends on two phenomena. The first one is a bibliometric regularity rooted in the main features of competition in the open space of science, that can be modeled by bibliometric laws, the parameters of which are "coverage-independent" indicators. But this regularity is obscured for many countries by a second phenomenon, the presence of a sub-population of journals that does not reflect the same international openness, the nationally-oriented journals. As a result indicators based on standard SCI or SCISearch perimeters are jeopardized to a certain extent by this sub-population which creates large irregularities. These irregularities often lead to an over-estimation of share and an under-estimation of the impact, for countries with national editorial tradition, while the impact of a few mainstream countries arguably benefits from the presence of this sub-population.

Introduction

Classic indicators, used to measure the scientific output of countries or institutions, are mainly based on ISI databases and in particular the *Science Citation Index* (*Garfield*, 1955). The particular properties of the SCI, such as its multidisciplinary coverage, the process of selection of journals, the cover-to-cover journal treatment, the full address identification and the reference recording explain its popularity for the building of science indicators, in particular those associated with the "mainstream" of internationalized science. The competition for access to journals brought about by the release of journal "impact factors", in ISI's *Journal Citation Report*, introduces a

Received September 16, 2002.

Address for correspondence:
MICHEL ZITT
Observatoire des Sciences et des Techniques (OST),
Paris (France)
E-mail: zitt@nantes.inra.fr

0138–9130/2003/US $ 20.00

positive feedback on the validity of this journal collection. However SCI perimeter is not sacred. ISI sources are available in a variety of versions, including the SCI, SCI-Search (a component of *Web of Science*), *Current Contents*, all candidates for bibliometric indicators calculation. Moreover a fairly significant turnover of journals is maintained in order to keep pace with the changing configuration of science. The current concern about bibliometric indicators and "benchmarking" in science policy raises more than ever the question of the robustness of measures – particularly those based on ISI sources – used to assess the outputs of scientific systems, especially national systems. A variety of methodological techniques, including, for instance, multiple author counting options, may explain apparent divergences between international science indicators published by specialized agencies (e.g., *Moed,* 1996).

Of the greatest significance also are the 'perimeter' attributes of the database (coverage, type of documents, etc.) which may dramatically alter the picture of world science. We focus here on the influence of what is perhaps the most important factor, the selection of the journal set.* One might expect that some regular bibliometric trend should be found when extending the dataset toward 'low level' journals, likely to influence the balance between leading countries and emerging countries; then, one might expect that the extension of the dataset would reduce statistical fluctuations in relative indicators such as countries' shares of world science. However we will demonstrate how this expected regularity is obscured by other phenomena of a large magnitude, which create a more complex situation for the observation of world science.

After an exploration of the data, methods and indicator definitions, we will provide empirical evidence of the significant changes in output series as a function of journal set extension. A specific irregularity and its major source are examined. After correction for this source of irregularity, a simple bibliometric approximation of the classic indicators' behavior is introduced. We summarize the changes in the picture of science with and without perimeter corrections. These changes which as we will see may be rather dramatic are discussed in the final section.

Data and methods

We used a customized extract of the *Integrated Citation File* (ICF) from ISI, at the Observatoire des Sciences et des Techniques (OST, Paris). The publication year 1997

---

* First stages of this work and provisional results were presented at the 6th *Int. Conference on Science and Technology Indicators*, Leiden, The Netherlands, 24-27 may 2000 and at the 8th *Int. Conference on Scientometrics and Informetrics*, UNSW, Sydney, 16-20 july 2001.

was chosen as a trade-off between timeliness and delay for citation accumulation. The complete dynamic set of ISI journals was addressed. This set is comprised of 3839 journals for the SCI set plus 1632 journals for the SCISearch extension (SCISearch is also the basis of *Web of Science* for natural sciences). The counts used are fractional ones, both for authoring, when a publication bears several institutional addresses, and for journal assignments, when a journal is assigned to several categories by ISI. Only articles, reviews, notes and letters are considered.* Citations are calculated at the paper level and the window of citation is 3-years (i.e., database years), including the year of entry. The same type of measure is used to calculate journal impacts starting from the document level. Differences in our definitions and the timeframe used account from slightly different impact ranks when a direct comparison is made with ISI's "impact factor" rankings.

Field discrepancies of citation behavior have been studied extensively since *Moravcsik* and *Murugesan*'s classic works (1978). The subject was reviewed by *Schubert* and *Braun* (1996). The optimum level of normalization (discipline, sub-disciplines, specialties...) remains a difficult question (see also *Kostoff*, 1997). Field-normalization mechanically improves the ranking of journals belonging to low-impact disciplines, such as mathematics, and conversely for journals belonging to high-impact disciplines, such as fundamental biology. Country profiles may be significantly different following normalization as a result of specialization patterns. Results reported here rely on field-normalized journal impacts, using the OST 8 disciplines + multidisciplinary scheme, which are based on the ISI sub-fields, whilst excluding social and human sciences. Analysis at other levels, such as at the level of specialty, may also show significant indicator distortion.

Impact is a natural criterion for journal selection. Assuming journals are ranked according to their decreasing (field-normalized) impact (j=1..n), a journal set J(r) is defined as the collection of the r journals with ranks $j \leq r$. In general terms, our focus is on the way the world scientific landscape changes with r. We pay particular attention to the range of large values of r (low-impact range), where the question of perimeter adjustment is very important from the practical point of view.

Another important feature of journals considered in the followings is their level of national orientation. We have limited ourselves to a very simple measure in this study.

* Two remarks on the set: first, Computer and Mathematics Citation Index (CMCI), in fact a thematic part of SCISearch, has been added to the SCI, as both are considered for standard indicators at OST. Then, as far as document types are concerned, articles from conference proceedings are excluded but their inclusion is open to debate (see *Sigogneau*, 1999).

This measure belongs to the family of relative indexes which compare the national authoring profile of journals (distribution of authors among countries) with the known profile of the discipline. This index is based on the maximum relative deviation* – this measure and more complete ones are discussed further in *Zitt* and *Bassecoulard* (1998, 1999).

The classic indicators under scrutiny include both the absolute ones (production volume, citations volume, impact of a country) and the relative ones (production world share, citation world share, relative impact of a country). Given that $p_i(r)$ is the publication of country i in the journal ranked r; $c_i(r)$ its observed citations; f(r) the impact (citations per publication) of the journal; $ec_i(r)$ the expected citation of country i in this journal, defined by $ec_i(r)=p_i(r)\times f(r)$, number of citations that would be received if each article had the average impact f(r) of the journal, then:

$P_i(r)=\Sigma_{j=1,r}p_i(j)$;
$EC_i(r) =\Sigma_{j=1,r}ec_i(j)$;
$EI_i(r)=EC_i(r)/P_i(r)$

denote respectively the publication volume, the expected citation volume, the expected impact of the country i for the set of journals through rank r, and

$C_i(r) =\Sigma_{j=1,r}c_i(j)$;
$I_i(r)=C_i(r)/P_i(r)$;
$RCR_i=C_i(r)/EC_i(r)$

denote respectively the observed citation volume, the observed impact and the relative citation ratio (*Schubert* and *Braun*, 1986) of the country i for the set of journals through rank r.

While $P_i(r)$, $EC_i(r)$, $EI_i(r)$ can be interpreted as the achievement of country i in inter-journal competition, RCR_i (r) describes their performance within journals. Finally, $C_i(r)$ and $I_i(r)$ relate to the overall competitive position of country i.

Relative indicators (publication share $PS_i(r)$, expected citation share $ECS_i(r)$, observed citation share $CS_i(r)$) are defined as ratios to respective world values (m countries).

---

* For a journal, the index is based on the authoring country which exhibits the maximum difference between its world share in the journal (s) and its world share in the discipline (m), the index being normalized by (1-m). This value is one of the simplest measure of national orientation of a journal. In a journal, the country for which the above deviation reaches its maximum is not necessarily the country of the commercial publisher, but the two coincide in most cases for strongly national oriented journals. It should be noted that absolute measures (number of countries, maximum share, concentration indexes...), that do not need decisions about discipline delineations, are unfortunately not satisfactory for journals with atypical patterns.

For example $PS_i(r)=P_i(r)/P_W(r)$, with $P_W(r)=\Sigma_{i=1,m}P_i(r)$. Similarly expected relative impact is defined as $ERI_i(r)=ECS_i(r)/PS_W(r)$, observed relative impact $RI_i=CS_i(r)/PS_W(r)$. Relative impact is a convenient synthetic indicator but has technical drawbacks in disciplinary aggregations.

A look at typical country series suggests that the behavior of major benchmarking indicators reflects the presence of both bibliometric regularities and significant perturbation factors. Figure 1 outlines the trends in relative indicators for the USA, Russia and India. In the first place it highlights the large interval of variation of classic relative indicators as r varies: the picture of science is strikingly different when comparing the small sets of highly visible journals and the complete SCI perimeter (sets inferior to 500 journals, yielding large statistical fluctuations, are not represented). It also shows sharp contrasts in variation between countries. The US world share of publications decreases rather consistently as r grows. Whilst on the other hand for India, an example of emerging country, we observe an increasing publication share when enlarging the set towards lower impact journals. The case of Russia is intermediary, but a striking point here is the break in trend in the tail of SCI, with a significant upsurge in publication share.

It follows from the definition of impacts that changes in citation volume, publication volume and impact for a particular country are intrinsically linked. For marginal changes, the citation variation (in relative terms) is the sum of the relative variations of impact and publication. This remains approximately true for discrete variations provided they are relatively small, but for countries with large variations in the interval under scrutiny, the second order term cannot be neglected (e.g. Russia). If citation volume is held constant in the interval, small variations of impact and publication volume, in relative terms, are symmetrical. This can also be extended to the corresponding relative indicators. On several occasions the approximate stability of citation shares allows us to observe this simple form of mutual dependence of publication share and relative impact.

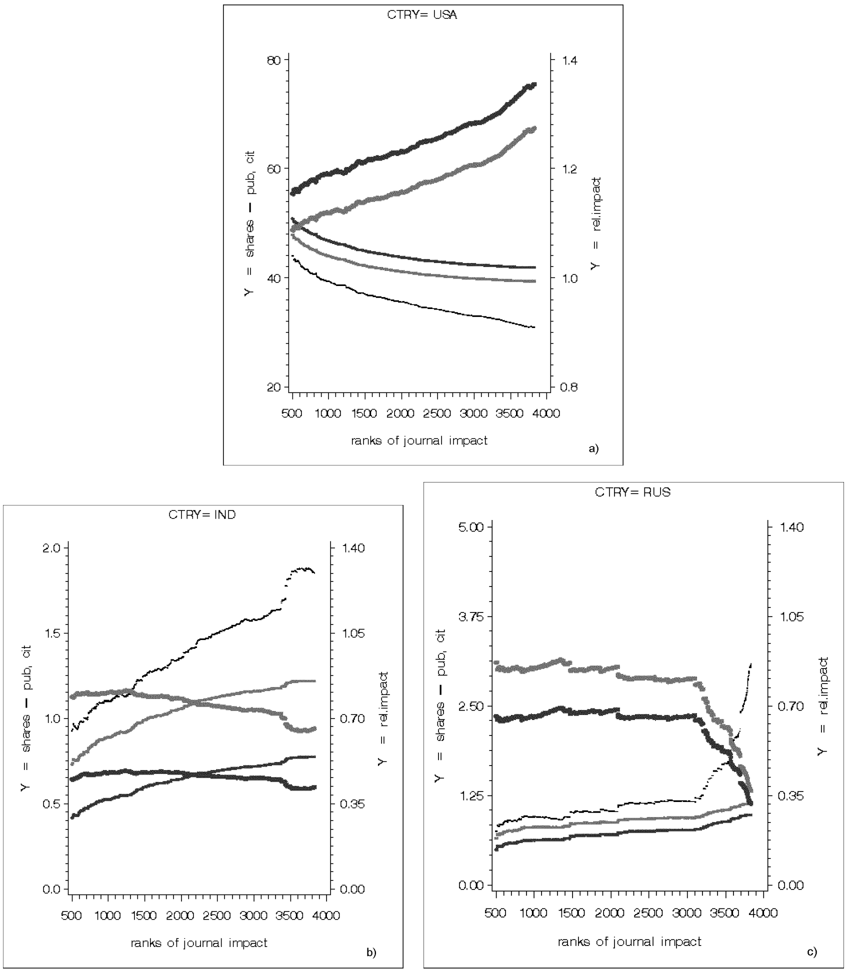

Figure 1. Standard science indicators as a function of database perimeter: original series (a. USA, b. India, c. Russia)

Thin line: world share of publication; pair of thick lines: relative impact (dark: actual; grey: expected); pair of medium lines: world share of citations (dark: actual values; grey: expected). USA, the leading country in science, unsurprisingly exhibits a decreasing trend in publication share (thin line, bottom) when the journal set extends. India increases its publication share (thin line, top) when the journal set expands, with an adverse effect on relative impact. A strong irregularity is observed in the right tail. Russia shows an intermediary trend, with a spectacular irregularity, an upsurge of publication share (thin line, middle) in the two last deciles of r and a concomitant collapse of relative impact. Actual and expected series for citations and impact are similar in shape, but their relative position is worth noting. The ratio observed/expected or "relative citation ratio" is largely above 1 for the US and below 1 for Russia and India.
All citation indicators are field-normalized.

Fundamental irregularities in the large perimeters range

A first observation, whatever the trend, is the amplitude of local fluctuations. The first differences series for country shares ($P_i(r)$-$P_i(r$-$1)$) records outlying deviations. The study of these deviations shows that they are far from being normally or log-normally distributed. In particular, for many countries, outliers exhibit large positive values, and also long sequences of very small (quasi-zero) negative values. The persistence and, in some cases, the increase in such fluctuations associated with cumulated indicators in the low-impact range (where the statistical basis is the largest) is not consistent with the hypothesis of an homogeneous population of journals. Large accidents in the series are likely to modify country share series according to a random staircase pattern. Figure 2 shows both publication series and their deviations for France and China. A first consequence of these deviations is that, even for models able to predict correctly the local trend across the range, the structure of errors will be strongly non-normal and unsatisfactory.

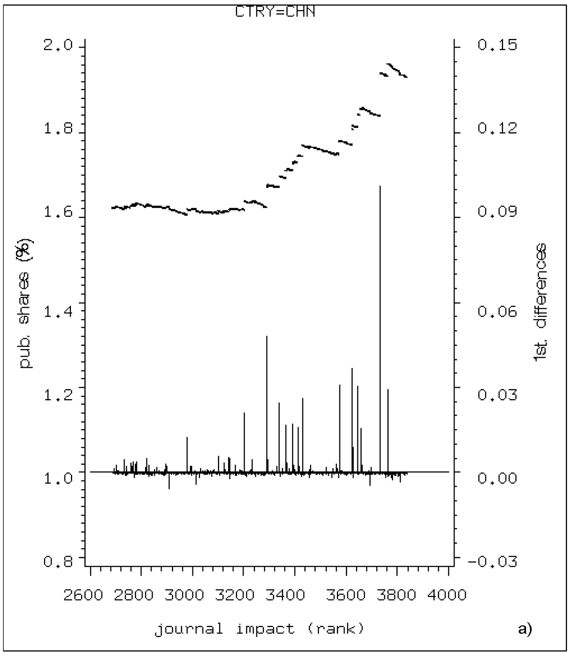

Figure. 2. Alteration of publication share series due to nationally-oriented journals, over the 3 last deciles of original series (a. China)
(For detailed explanation see Figure 2b on next page)

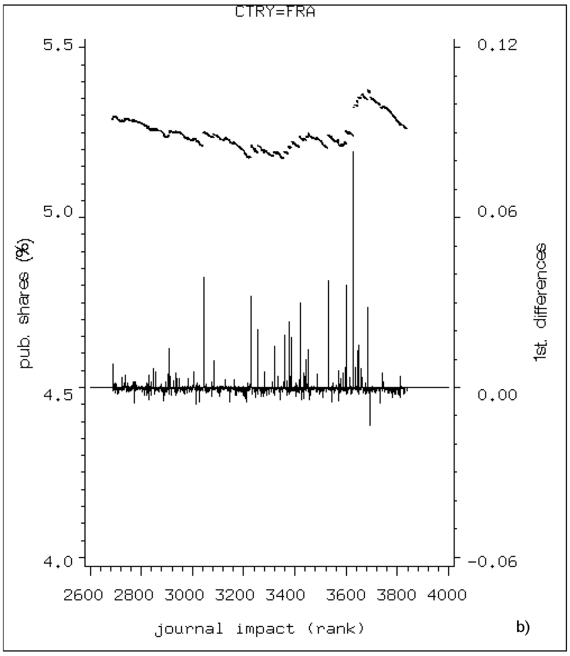

Figure. 2. Alteration of publication share series due to nationally-oriented journals, over the 3 last deciles of original series (b. France)
The upper curve shows the publication share series, the lower the corresponding first differences series (enlarged scale). For China peaks represent, by decreasing amplitude: *China Science Bulletin, China Physics Letters, Acta Chimica Sinica, Communications in Theoretical Physics, Science in China Series A, Chinese Medical Journal, Applied Mathematics and Mechanics*. For France, major deviations are recorded for *Presse Médicale, Comptes Rendus de l'Académie des Sciences IIA, Arch. de Pédiatrie, Arch. des Maladies du Cœur et des Vaisseaux, Gastroentérologie Clinique et Biologique, Comptes Rendus de l'Académie des Sciences IIB*

In some cases, the accumulation of irregularities shapes a new trend. The rise of publication share is spectacular for Russian series – see the right-hand tail (Figure 1). A similar feature is recorded for the Ukraine. No other country registers a doubling of publications or more within a range of two or three deciles of impact. In contrast, breakpoints in trends, with a downward tendency, are detected for several countries (e.g., Nordic countries, Italy). These behavioral changes for particular countries are echoed in the global structure of world publications shares, that also undergo unexpected moves in the low-impact range.

These observations can be connected with a unique source of irregularity, namely the presence of outliers with large positive values in country series, which cause strong fluctuations or, when accumulated in the same range of r, breaks in trend (Russia). These outliers prove to be associated with the nationally-oriented journals (NOJs) detected with the aforementioned measures of internationalization. The distribution of journals on the internationalization index has a long tail or a bi-modality, according to the discipline, caused by the minority of NOJs (*Zitt* and *Bassecoulard*, 1998). In a situation of bibliometric regularity, journals may be considered as samples of world production at a given "prestige" level (in the neighborhood of some r value). In other words their authoring profile statistically reflects the international balance in the discipline at this level. To be able to disclose this regular behavior, we had to identify and remove the source of massive irregularities. A few NOJs are big enough to alter publication shares even based on large datasets. This situation of not infrequent remote outliers, whose individual contributions are able to modify measures based on large cumulations, is typical of heavy tail distributions.

The proportion of NOJs greatly varies significantly with the impact rank r. If for convenience we consider as nationally-oriented the 20% journals with lower internationalization index in the particular period studied (a reasonable choice considering the internationalization index distribution), the 8th decile of impact is slightly above the expectation (21.4% of journals are NOJs), with the 9th decile at 34.1%, and the 10th at 56.2%. The figures are even greater for SCISearch. This phenomenon partly accounts for a certain amount of correlation between impact and internationalization measures, which is low to moderate at the discipline level. It should be noted that, the distribution of NOJs among countries is most uneven. The distribution of national origin of these journals (Table 1) does not tally with the scientific output among the countries. Germany, France, India, China and especially Russia are over-represented with respect to their share of output in science. Conversely, the accumulation of NOJs from particular countries tends to dampen the representation of other countries that, whatever their scientific background, have a weak nationally-oriented editorial tradition as far as SCI-covered journals are concerned. In the first difference series, this accounts for sequences of negative quasi-zero values in the range of large r. Examples are the Nordic countries and the Netherlands, the latter with a strong editorial power but internationally minded. Even though the internationalization of journals has been steadily increasing in the last few decades, the proportion of NOJs remains a serious source of irregularity.

Table 1. Nationally-oriented journals in the low-impact range (1997). Distribution according to the dominating country*. Ranked by percentage of publications in the range (SCI)

Dominating country	Number of journals (SCI) (%)	Publications in these journals (SCI) (%)	Number of journals (SCI expanded) (%)	Publications in these journals (SCI expanded) (%)
RUS	16.3	27.0	10.1	19.4
USA	21.2	14.3	26.1	17.8
GBR	7.9	9.7	5.8	6.2
DEU	11.0	8.5	10.1	9.1
JPN	9.1	7.4	8.6	9.1
FRA	6.3	6.5	6.9	7.2
CHN	3.0	4.7	3.1	4.9
AUS	3.0	3.3	2.1	2.4
IND	3.1	2.9	4.2	5.2
CAN	2.8	2.4	2.0	2.0
…				
World	100.0	100.0	100.0	100.0

* In a nationally-oriented journal profile, the country that records the maximum deviation stated above (often the publisher's country). The low-impact range is defined as the last 3 deciles on r, in SCI. The proportions are sensitive to the choice of a particular internationalization index.

Examples of the strong impact of particular journals on series can be found in many countries' profiles:

Peak for China and France are commented in Figure 2. Other examples are, for Germany, *Deutsche Medizinische Wochenshrift, ROFO-Fortschritte*. The national orientation is not only a matter of language, with in India *Current Science*, the *Indian Journal of Chemistry A/B and Pramana-Journal of Physics* having significant influences. For Spain, a strong singularity is created by *Medicina Clinica*. Russia offers a wide collection of NOJs, the most remarkable being *Doklady Akademii Nauk* (Academy of Science) which carries more than 1300 articles and boosts the Russian share, with a noticeable negative effect on other countries' series, and we also find effects for the *Russian Chemical Bulletin, Zhurnal Fizicheskoi Khimii, Zhurnal Obshchei Khimii, Physics of the Solid State*, etc. The chaotic aspect of some country series may impede a proper identification of trends. However, if most fluctuations due to national journals occur in the last deciles, there are a few exceptions, a prime example being the *British Medical Journal*, the strong impact of which shapes the UK publication series with an early staircase deviation. It can be seen that many deviations are due to medical research journals, a stronghold of national traditions.

As a result, the structure of the world publications is much less stable in the "real" SCI perimeter than might be expected. Even the citation shares, the most stable indicator, reflect slight irregularities (partly as a result of the Russian and Ukrainian upward trends). In order to examine the bibliometric regularity, and to assess the variations of classic indicators in several hypotheses, we built, along with the original series, a corrected series where rules for discarding NOJs were applied. Obviously, only the NOJs with low-impact should be removed, since a minority of the "home-grown" journals can be found in the medium impact range, as noted above for the BMJ. A few are even found in the very top impact range, namely journals dominated by US authors (one should remember that the US share, reflected in the model above, peaks in this region). This led us to suggest a possible modulation of internationalization measure by the level of impact (*Zitt* et al., 1998). Apart from "review" journals with often so few articles that the significance of the internationalization measure is questionable, one also finds the case of high-standard journals in particular areas dominated by a single leading country, or else emerging journals that start with a "national phase" before internationalization (*Leydesdorff* and *Cozzens*, 1993). Here we chose to remove, for our corrected series, nationaly-oriented journals (deciles of internationalization 9-10) belonging to the deciles of impact (field-normalized) 8-10. One of us used a slightly different option in a first attempt at perimeter adjustment for the OST indicators (in *Barré* et al., 1999).

Let us now consider the expanded version of SCI, SCISearch, which is the basis of the Web of Science (for natural sciences). We calculated impacts for these additional journals normalized on the SCI field-averages. This extension would be expected to extrapolate the structure of the last deciles of SCI, with a high proportion of irregular journals. The observed picture is quite different. SCISearch actually brings in its own set of NOJs, but also regular ones that deserve a rather better ranking than the tail of the SCI on both normalized impact and internationalization criteria. Using the same absolute values of thresholds that led us to discard 420 journals from the SCI in order to build the corrected series, led us to eliminate 627 journals out of the 1632 representing the SCISearch extension with respect to our set SCI/CMCI. In other words a thousand journals from the extension would be ranked better than the tail of SCI. In rare cases, high-impact journals do not yet belong to the SCI, perhaps being a provisional position, these journals being introduced later to the core. In some cases, methodological choices can also interfere (in our impact ranking, proceedings are not counted; ISI selection criteria are not only based on impact, we bring CMCI to SCI together, etc.). As a result, a journal is generally given different rankings in the two series SCI and SCISearch, hence the country series for a given indicator, say for publication share, can slightly diverge for lower values of r. The discrepancies are commented upon below.

Regularities: A simple bibliometric approximation

The evolution of countries' indicators as a function of the impact-based journal set is expected to reflect fundamental features of publication and competition processes in science. For example, leading countries are likely to be relatively more present in a small collection of high-impact journals (low r), and less present in a collection extended through low-impact ones (high r). Furthermore, the stability of countries' indicators, given the principle of ranking, is likely to be better for citation indicators than for publication indicators in the low-impact range (high values of r), because of the higher concentration of citations. This has to be systematically examined.

A classic choice for empirical modeling of concentration phenomena is the use of hyperbolic functions. Power laws first used in economics (Pareto distribution) have been successfully applied to several categories of informetric problems (typically Lotka's productivity law, Bradford's law of bibliographic scattering, Zipf-Mandelbrot's law on word distribution) as reviewed by *Haitun* (1982) or *Egghe* (1991). As well as classic interpretation of a cumulative advantages process (*Price*, 1976), the power-law encompasses several classes of causal models in bibliometrics (*Bookstein*, 1990), with more recently a particular emphasis on self-similarity, following the development of fractal theory. A direct application in the country indicators context is found in *Katz* (1999), who studied the scale invariance of the citation-publication relationship. Pure concentration curves, for example cumulated citations as a function of journal citation ranks, can be described with the standard hyperbolic model. As $EC_w(r)$ or $C_w(r)$ are studied instead as functions of journal impact rank, lower but still strong concentration levels are expected (the two cases would only be identical if all journals had the same size). The publication volume $P_w(r)$ can be expressed with similar forms, with much lower concentration, expressing the relationship between the size and impact of journals.

The empirical model is based on the following approximations at the journal level: $p_i(r)= b_i\ r^{-\beta i}$; $ec_i(r)=g_i\ r^{-\gamma i}$; $c_i(r)=d_i\ r^{-\delta i}$, where the parameters are specific to each country i. The model can be fitted to the integral forms $P_i(r)$, $EC_i(r)$, $C_i(r)$ that represent the indicators of volume for the set of journals having ranks $j \leq r$.

For example $P_i(r)=[\int_{j=1,r} p_i(j)dj\ ] + b_i$ where b_i stands for $p_i(1)$.

Hence $P_i(r)= (b_i /(1- \beta_i)) (r^{(1- \beta i)}- \beta_i)$.

At the world level, the citation volume $C_W(r)$ (identical to $EC_W(r)$) and the publication volume $P_W(r)$ are similarly obtained from c_W and p_W. The relative indicator, for example publication share, is given by $PS_i(r)=P_i(r)/P_W(r)$. In the general case the expression is rather intractable for direct fitting purposes. However for a large "r" and empirical β values <<1, the quotient may be approximated by the power law of exponent $\beta_W - \beta_i$:

$PS_i(r)=[\ b_i(1-\beta_M)\ /\ (\ b_M(1-\beta_i\ )\ )\ ]\ r^{(\beta_W-\beta_i)}$

For the reasons mentioned above, the bibliometric approximation can only be operationalized on the original series where last deciles are cut off, or on the corrected series without truncation (results are reported using the latter option). For fitting we also discarded the first two deciles, which were subject to large statistical fluctuations. The reported results are based on the simpler method: linear fitting after log transformation. β_i was estimated both directly on volume and indirectly on share series (β_W-β_i). The estimates were practically identical. R-Squares were excellent (> 0.98, mostly > 0.99) for volumes in all cases examined (29 countries). For shares, the goodness-of-fit was also high, except for countries with nearly flat trends, i.e. β_i close to the world value β_W. Non-linear fitting, with less leverage effects was also conducted, with little change in country ranking.

The USA and Switzerland are the only countries exhibiting a parameter β_i much higher than the world average β_W=0.31, namely ca. 0.46. As for these countries (β_W-β_i) is negative, their shares in publications monotonously decrease as r increases. Though their magnitude diminishes with r, the changes remain significant in the tail.

The next group of countries exhibits a lower dependence of publication share on r (β_i approaching the world value). France, Germany and the Netherlands are about β=0.34. The UK, Israel and Canada, exhibit a β practically at the world average, 0.30. The parameter for the UK would be higher without the singularity of the *British Medical Journal*. Between β=0.29 and β=0.23, corresponding to a slight upwards trend for shares when r increases, we find Italy, Belgium, Denmark, Spain and Sweden.

For all remaining countries, 0<(β_W-β_i)<1, indicating a marked upward trend of publication shares. Japan, Australia and Finland are at β=0.17. Next come countries with β scattered between 0.13 and zero: Korea, Russia, Taiwan, Norway, New-Zealand, Brazil, Poland. Czech Republic, Argentina, Mexico are slightly below 0, between -0.02 and -0.03, India and China are at β=-0.05. These countries show a negative value of the publication parameter, implying than the higher r, the larger the world share and also the larger the absolute number of publications in journals of rank r.

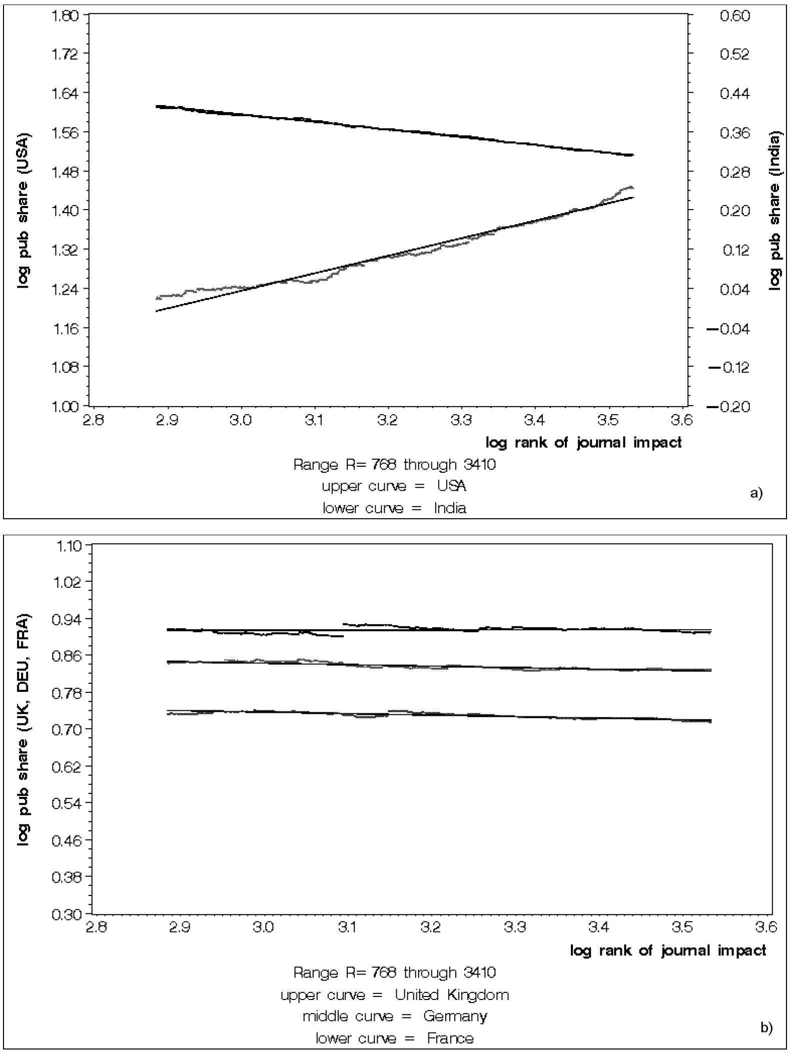

Figure. 3. Publication shares as a function of database perimeter: bibliometric trends on corrected series, double-log scale (a. USA and India, b. Germany, France, UK)
The parameters are fitted on following approximations $P_i(r)=a_i\, r^{(1-\beta i)}$; $P_w(r)=a_w r^{(1-\beta w)}$ respectively for publication volumes of country i and the world w, $PS_i(r)=A_i\, r^{(\beta w-\beta i)}$ approximates publication shares of country i, with $\beta_W= 0.31$. The graph (a) shows the contrast between USA share series, β_i=0.46, and India (β_i =-0.05), using different scales. The graph (b) shows the quasi-stability of shares (very slight down-trend) for 3 large European countries (Germany, France, β_i =0.34, plus UK with the visible singularity of *British Medical Journal*)

Fitted publication shares are shown in the double-log plot of Figure 3 for a few countries in each group. The abcissa is the logarithm of the journal impact rank, whilst the ordinate is the logarithm of the publication share.

As expected, for all countries the citation parameters show more concentrated patterns than the publication parameter, with γ_i and δ_i much higher than β_i. For example γ is above 0.80 for the USA and 0.68 for Switzerland with decreasing world shares. Most European countries range between 0.72 (France, and also the world average) and 0.65 (Spain). Japan, Sweden, Finland and Austria have γ above 0.60. Next come Russia, Korea, Norway and Thailand. The lowest values (<0.45) are observed for Brazil (0.45), India, the Czech Republic, Mexico and China (0.37), with strongly increasing world shares. The observed R-Square is excellent for all countries.

γ_i and δ_i may be different by a few points but yield pretty close country rankings. The form and parameters of the model for a country's publication share already predict decreasing changes of this indicator as r increases, however with significant variations in the range for large r, particularly for countries deviating from the world average (such as US or peripheral countries). For citation shares, the stabilization occurs more rapidly and in the range of large r the predicted world structure is more stable.

It follows from the expression of publications and citations volumes and shares that, with respect to approximation conditions, relative impact on the one hand and RCR on the other (not shown here) can be modeled with power-law forms for sufficiently large r. The parameters of the law, or of an other appropriate model, can be considered as coverage-independent indicators for a comprehensive set of regular journals. As mentioned above, SCI missing a few good impact journals, the series of SCI and SCISearch do not coincide since ranking of journals diverge even for low r values. This moderate divergence, as well as the larger number of observations in the series, yield slightly different values of the power-law parameter.

Consequences for countries benchmarking

Variation of classic indicators as a function of the perimeter depends on regularities amenable to bibliometric modeling, often altered by large amplitude fluctuations due to NOJs. To what extent do the two phenomena influence the general picture of world science? The effects can be sketched either by letting r vary respectively on original and corrected series, or comparing original and corrected series for a given perimeter. Table 2 yields some results for a few selected countries, ranked by the relative variation within the original series (smoothing options are different from previously reported results on a former year):

- Column A gives an account of what a cut-off of the SCI based on journal impact yields.
- The sensitivity of the original SCI series to perimeter variations (observed), which involve both "natural" trends and irregularities, is shown in column B. We chose the interval of the 3 last deciles (r_1=beginning, r_2=end) in order to get a more stable basis for r_1, but in fact major irregularities occur in the last two. Indicators' values have been smoothed on the (arbitrary) basis of 50 journals (r, r-49) in order to reduce local fluctuations. The underlying rationale is a cut-off strategy discarding a low-impact sub-range.
- The sensitivity of the predicted series (power-law model on corrected series) to perimeter variations is shown in column C. This sensitivity is completely described by the parameters (for convenience the same interval as for column B is chosen).
- T he divergence of the original and corrected series (column D), for their respective complete perimeters. The differences mainly reflect the role of NOJ irregularity. The underlying rationale is a cut-off strategy that does not discard all low-impact journals, but only those with a strong national-orientation. Other comparisons might involve, for example, the original and the corrected series for the same number of journals, by cutting off at the maximum perimeter of the corrected series.

The original series of classic indicators, either SCI or SCISearch, is most unstable in the range of large perimeters. First let us bluntly adjust the perimeter by removing journals in the last deciles (perimeter 1). This has dramatic consequences for the comparative assessment of Russian vs. western science, or US vs. emerging countries. The effect is most apparent for Russia's share (especially in the two last deciles). A look at columns C and D suggests that this spectacular variation can be mostly attributed to NOJs. To a much lesser extent, an important increase in the last three deciles of the original series is recorded for India and China, partly due to the presence of NOJs. Poland and Brazil also gain publication shares, but mainly because of the trend. With lesser editorial power than the former, they don't benefit from the addition of NOJs to the series. New Zealand, Chile, Czech Republic and Argentina, not shown, also gain some points in the last deciles.

Table 2. Variation of publication share indicators (SCI/CMCI)

	A	B	C	D
Country	Restricted perimeter original series v_1*	Difference deciles 8-10 original series v_2*-v_1*; (v_2*-v_1*)/v_1* (%)	Difference deciles 8-10 predicted series v_2**-v_1**; (v_2**-v_1**)/v_1** (%)	Difference between maximum perimeters of original and corrected series (v_2*-v_3*); (v_2*-v_3*)/v_3* (%)
NLD	2.26	-0.19 (-8.5%)	-0.02 (-1%)	-0.14(-6.4%)
USA	36.60	-2.84 (-8.4%)	-1.84 (-5.4%)	-1.59(-4.9%)
SWE	1.71	-0.09 (-5.1%)	0.05 (+2.9%)	-0.13(-7.5%)
FIN	0.80	-0.04 (-5.4%)	0.04 (+5.2%)	-0.06(-7.3%)
CAN	3.97	-0.16 (-4.0%)	0.01 (+0.3%)	-0.08(-2.1%)
JPN	8.97	-0.36 (-4.0%)	0.44 (+5.0%)	-0.16(-1.9%)
…				
GBR	8.27	-0.14 (-1.7%)	0.01 (+0.1%)	-0.00 (-0.0%)
FRA	5.30	-0.02 (-0.4%)	-0.06 (-1.2%)	0.07 (+1.3%)
DEU	6.72	0.15 (+2.2%)	0.08 (+1.1%)	0.12(+1.8%)
…				
BRA	0.73	0.04 (+5.3%)	0.07 (+9.9%)	-0.01(-1.9%)
KOR	0.93	0.08 (+8.6%)	0.06 (+6.8%)	0.05(+5.7%)
POL	0.66	0.07 (+10.1%)	0.08 (+11.2%)	-0.00(-0.4%)
CHN	1.62	0.32 (+19.9%)	0.22 (+13.7%)	0.26(+15.4%)
IND	1.53	0.33 (+21.5%)	0.21 (+13.6%)	0.10(+5.5%)
RUS	1.17	1.81 (+155%)	0.09 (+7.7%)	1.79(+152.0%)

SUBSCRIPT :

v_1: refers to the perimeter defined by the beginning of the 8th decile, r=2688, common starting point for original and corrected series.

v_2 refers to the complete SCI/CMCI non corrected, r=3839

v_3 refers to the complete SCI/CMCI without NOJs of deciles 8-10, i.e. corrected series, r=3410

Values: * smoothed observed value; ** predicted (power-law model).

Ex: v_2** denotes the predicted (extrapolated) value for r=3839, v_3* the observed value for r=3410 on corrected series

Values are sensitive to methodological options (documents types, type of normalization, etc.)

In contrast, the USA and the Netherlands show a strong decrease in publication share (more than 8%) followed by Sweden, Finland and Canada (around 5%). Japan, Belgium, Norway, Israel and Denmark also join this group of countries with a clear decrease in publication shares. Sweden and Finland particularly illustrate the case of an upsurge in publication share when the NOJs are removed. Major European countries exhibit lower levels of change. Germany and France show a slight increase which may be related to the presence of their collection of NOJs. Output measures of large European countries and Israel are the least sensitive to perimeter variations, especially in corrected series.

The effect on highly correlated indicators such as "expected citation shares" and "expected relative impact" is automatically strong for countries in extreme positions, leaders and emerging countries (not shown), but real citations and impacts behave by and large in the same way as their expected counterparts. Table 3 shows the variations of observed citations and relative impact, in the original series for the same countries as above.

Table 3. Variation of citation shares and relative impact indicators (SCI), 1997

Country	Difference on deciles 8-10 original series v_2*-v_1*; $(v_2*-v_1*)/v_1*$ (%)		Difference between maximum perimeters of original and corrected series (v_2*-v_3*); $(v_2*-v_3*)/v_3*$ (%)	
Rationale	Truncation of SCI on journal impact criterion		Truncation of SCI on journal internationalization and impact criteria	
	Citation shares	Points of relative impact (world=100)	Citation shares	Points of relative impact (world=100)
USA	-0.9 (-2.0%)	8.7 (+6.9%)	-0.39 (-0.9%)	5.4 (+4.2%)
NLD	-0.02 (-1.0%)	8.6 (+8.1%)	-0.03 (-0.9%)	6.4 (+5.9%)
SWE	0.02 (+1.1%)	6.2 (+6.5%)	-0.02 (-1.1%)	6.5 (+6.9%)
FIN	0.00 (+0.5%)	5.9 (+6.2%)	-0.01 (-0.9%)	6.5 (+6.9%)
CAN	0.00 (+0.1%)	4.0 (+4.3%)	-0.01 (-0.2%)	2.3 (+2.4%)
JPN	0.01 (+0.1%)	3.5 (+4.3%)	0.02 (+0.3%)	1.8 (+2.2%)
…				
GBR	-0.00 (-0.2%)	1.6 (+1.6%)	0.01 (+0.1%)	0.1 (+0.1%)
FRA	0.00 (0.1%)	0.5 (+0.5%)	0.01(+0.2%)	-1.1 (-1.1%)
DEU	0.07 (1.0%)	-1.2 (-1.1%)	0.04 (+0.5%)	-1.3 (-1.2%)
BRA	0.02 (+4.3%)	-0.5 (-0.9%)	0.00 (+1.1%)	1.5 (+3.0%)
KOR	0.03 (+5.4%)	-1.8 (-3.0%)	0.02 (+3.9%)	-1.0 (-1.7%)
POL	0.03 (+7.5%)	-1.4 (-2.4%)	0.01 (+2.4%)	1.5 (+2.8%)
CHN	0.06 (+7.5%)	-5.4 (-10.3%)	0.04 (+4.3%)	-5.0 (-9.6%)
IND	0.07 (+10.1%)	-4.3 (-9.4%)	0.03 (+3.8%)	-0.7 (-1.6%)
RUS	0.21 (+28.7%)	-32.4 (-49.5%)	0.21 (+26.9%)	-32.4 (-49.6%)

subscripts, see Table 2

As a result of a stronger concentrated scheme, citation shares are less sensitive to the extension of the set than publication shares in the higher ranges of r. However this stability, which was expected from the principle of set construction – based on the impact criterion for journal inclusion – is not complete. Some changes are recorded, especially for Eastern European and emerging countries, which gain significant citation shares in the last deciles (original series), although to a lesser extent than publication shares. As a result the relative impact is down, with Russia once more as the dramatic example. In the 3 last deciles of the original series, the USA and the Netherlands, gain more than 8 points of impact, Sweden and Finland more than 6 points and Canada and Japan between 3 and 4 points. For Nordic countries, this is predominantly due to the effect of other countries' NOJs.

At the other extreme, keeping the three last deciles of the original series yields a collapse of Russia impact by 32 points, principally due to NOJs. To a much lesser degree (5 points, almost 10% of the impact value) this is also true for China. For Brazil or Poland the decrease of impact in the original series cannot be explained by the NOJs irregularity.

Among European countries, France and Germany record a significant reduction in impact with the addition of NOJs. Italy and Belgium (and also Israel) gather impact points due to the relative scarcity of their own NOJs.

A second source of indicator divergence lies in the extension of the database. We have mentioned that the extended SCI (SCISearch) collects a fairly large quantity of journals in a better position (both in terms of impact and internationalization) than those belonging to the tail of SCI. So the irregular tail of SCISearch is displaced towards higher values of r.

To give a first idea of the divergence, we can at first compare values of publication share for the respective complete perimeters of the SCI and SCISearch (original series). India (23% gain on the publication share) and Brazil (15%) are the most sensitive to the extension, followed by Poland, Russia (>10%) and China. Accordingly, India and Brazil lose more than 5 points in relative impact i.e. more than 10% of their relative impact value. On the other hand, Denmark, the Netherlands, Israel, Sweden and Switzerland lose between 4 and 7% in publication share. These countries however gain between 4 and 7 points of impact. These differences embody the effects of a bibliometric trend as well as of the NOJs irregularity.

We can try to neutralize the influence of NOJs by removing journals at the absolute impact and internationalization thresholds already practised for building the corrected SCI series. The respective corrected perimeters gather 3410 journals (SCI) and 4384 (SCISearch). The divergence between country publication shares in the two databases appears as much smaller when comparing their respective corrected series. Mexico and India gain more than 5% publication share in SCISearch. Among countries with lower publication shares in the extension, Japan looses almost 4%, next come China and France (>2%). In terms of relative impact, the differences in corrected series are low, only Mexico (-2.2) is slightly outside the range [±2 points].

Discussion and conclusion

Bibliometric indicators are methodology dependent. The counting options, especially for citation measures, and the type of normalization, may cause serious divergences in country output assessment. Here we have focused on another major source of sensitivity, the extension of datasets within ISI coverage, when the datasets are defined as collections of journals based on impact ranking. Country indicators, especially publication and impact, undergo variations when the journal dataset is extended toward lower impact items. This is firstly due to a "natural" trend which is predictable through bibliometric modeling that put leading (US, Switzerland etc.) and emerging ones (e.g. Brazil, China, India) in opposition to one another, with large European countries in the middle. Practical results for international benchmarking are easily deduced.

This trend has been approximated after correcting for the main source of irregularity, the NOJs. Keeping them in the dataset tends to push up the publication figures of countries with a strong nationally-oriented editorial power (whatever their position, second-best, intermediate or emerging) and to bring down values of other countries. The effect on relative impact, another classic benchmarking indicator, is also significant: the original perimeter of the SCI dramatically under-estimates the impact of the Russian literature of international standing, and moderately that of emerging countries with national editorial power such as China and India. These results suggest that the biases in the SCI cannot be reduced to the disputed issue of over-representation of US literature. Over-representation of a particular kind of literature of other countries, with effects on impact, may also matter.

Several points about our results need to be discussed. As far as the bibliometric approximation is concerned, the model is over-simplified, and residual trends appear in some cases. The model may be improved in a number of ways.* Alternatives to power-laws have been proposed for inter and intra-journal competition (*Van Raan*, 2001).

* A more elaborated form for C_i can be deduced from $p_i(r)$ and $f(r)$, using $C_i(r) = \int_{j=1,r}(p_i(j) \times f(j))$. The resulting form is less adapted for fitting purposes, but expresses the fact that the knowledge of journal impact distribution and country publications are sufficient to describe a first set of classic indicators (publication and expected citations/impact). Modelling of RCR is necessary to deduce real citations estimation. However, a correction for high values of r, for example $p_i(r) = (c_i\, r^{-\beta i}) - d_i$, can also be practised for journal impact.

It remains true that the power-law approximation is satisfactory in most cases, and addresses the range of classical indicators in a flexible way. Other types of indicator, for example Matthew effect at the country level (*Bonitz* et al., 1999), are also intimately linked to concentration phenomena.

The results will need to be extended to more recent years (with the constraint of a sufficient citation window). Other methodological choices, such as the options of field-normalization of journal impacts also influence the trend of indicators. As mentioned earlier, the choice of a good normalization level is a recurrent difficulty in bibliometrics. As far as type of document is concerned, "proceedings" could also be considered. The options taken for measuring impact, or internationalization may also affect the results. For the measures of internationalization, the profile of science used for reference may be defined in a more sophisticated way, taking into account the level of impact and an iterative process to neutralize the effect of national-journals in the reference profile. The construction of the corrected series, based on observation of distributions, can also be refined.

However, adjustment of settings within sensible limits is not likely to alter the major findings of this study. National benchmarking indicators are sensitive to the delineation of perimeters, so which database (or version thereof) is more appropriate for the construction of international benchmarking indicators? It has often been said that the SCI should be largely expanded (*Moravcsik*'s argument in the Philadelphia controversy: *Moravcsik*, 1988) to provide a better coverage of science in emerging countries. Dynamic adjustments to SCI coverage periodically raises questions (*Basu*, 1998). The role of the nationaly-oriented journals for the dissemination of knowledge cannot be denied. On the other hand, as far as bibliometrics is concerned, there is little doubt that the inclusion of journals with very low levels of impact in a benchmarking set is questionable.

A first argument is the low reliability of the impact measure when a low impact is associated with a low internationalization of journals, a small size and sometimes irregular publications. The impact measure for low-impact journals of small size relies on very few statistical events (citations) and is likely to be most unstable. Then the impact measure of NOJs may be jeopardized by particular referencing habits of a small community, especially when citations come from the same country (the internationalization of citation sources can be further measured in the same way as authoring, as we have suggested in the papers cited above). At low level of impact and internationalization, differences in impacts may be of little help to establish a hierarchy of journals from various national communities (say a Russian, a French and a Indian journal), so that the decision of inclusion or rejection of the journal is largely arbitrary.

Secondly, we have shown that the inclusion of low-impact and low-internationalized journals jeopardises the regularity of classic science indicators, which are based on the assumption of an internationally competitive science space. Adding a population with other norms and functions (such as transfer of knowledge, or technological content – see below) may be discussed but anyway introduces a fundamental heterogeneity. Narin, at CHI Research, has already conducted some selection for NSF indicators, in the context of a constant journal set. A radical proposal for SCI 'cut-off' was put forward by *Sivertsen* (1992). A possible restriction of SCI has also been suggested on the grounds of language biases (*Van Leeuwen* et al., 2000). Language is a clear marker of national-orientation, but origin of institutions leads to a more general qualification.*

As far as countries with a national editorial tradition are concerned, and in particular Russia and emerging countries, the restriction envisioned here would not necessarily tarnish their image. "A large publication share with a low impact" can in some contexts give a poorer picture of a national output than "a smaller publication share, with a better impact performance". But more subtle situations are found where peripheral journals also bring citation gains, as seen for India. The complex citation phenomena in peripheral groups of journals deserve further study. As for many mainstream countries, their impact rating largely benefits from the questionable presence of a sub-population of irregular journals. A corrected set tends to give a more balanced perspective.

The idea that the international science space is the norm may seem heavy in some instances. Russia inherits a situation, where some high-quality journals could only gather citations from within, resulting in a severe under-scoring. In other countries, national reference media with little international circulation, such as National Academy of Science journals, outside the mainstream, also deserve further examination.

Another point we made is that even whilst maintaining a high standard of selection, the issue is not only "restriction" of the SCI. Investigating the "extended SCI coverage" (SCISearch), we found that a fairly large proportion of journals in the extension perform better in terms of impact (and internationalization) than the tail of SCI as such. In any respect, building a good dataset for bibliometric benchmarking, at the country scale but also at other scales relevant for micro-bibliometrics, has to consider the continuity between SCI and more extended databases, in addition to other classic issues (types of document etc.). In a broader context**, not only SCI but all databases used for information retrieval or bibliometrics face coverage and usage problems. A main issue is the coverage of the technical literature, for which academic impact may not be a

* In a dynamic view we studied the disconnection between language/ origin of authors/ origin of publishers as a sign of transition of publication systems towards the trans-national model of science (*Zitt* et al., 1998).

** We are indebted to an anonymous referee for these important points.

better criterion than linkages to applications. The "fundamental" bias of SCI has been often stressed and recently *Kostoff* (1998) proposed a diagnostis of coverage and usage limits of this database, especially with respect to technology-oriented science. The multiple functions of knowledge, addressed in "Mode-2" debates, pose questions to the scientific information system.

Back to the narrower point of view of the academic science, the combination of a corrected series and high quality modelling would make useless the quest for the "optimal cut-off" in ISI coverage, as a few bibliometric parameters would globally describe the publication behavior of every country, over a wide range journal collections. For this purpose, the flexibility and simplicity of power-law approximations is an advantage, but other forms may be more precise. It is clear however that figures corresponding to one or two levels of cut-off are more acceptable in a science policy context than synthetic bibliometric parameters. The availability of straightforward indicators such as "market share" or "impact" is an important issue in the debate on science policy in the evolving context of science and society relations (*Gibbons*, 2001, *Barré*, 1999). In this respect, the fine tuning of corrections/restrictions within sensible limits is necessary (given an extensive study of impact and internationalization distributions). The stakes are high for macro-level (international comparisons) as well as micro-level studies (institutions strategic positioning). Finally, an unavoidable issue in this context is the current status of scholarly journals in scientific communication. How long will the scientific journal remain the major channel of communication for most disciplines? Threats to the functions of journals come from the electronic revolution on the one hand, and the pressures for proprietary forms prevalent in applications of science, such as patents, on the other. As long as peer-reviewed journals are a good tool for scientific communication, and hence for description of the international scientific landscape in most disciplines, it should be remembered that rather different pictures of world science may be drawn depending on the choice of database and its perimeter of inclusion.

References

BARRÉ, R. (1999), Ed., *Science & Technologie, Indicateurs Edition 2000*, Rapport de l'OST, dec. 1999, 464 p, Economica, Paris.

BARRÉ, R. (2001), The agora model of innovation systems: S&T indicators for a democratic knowledge society, *Research Evaluation* 10 (1) : 13–18.

BASU, A. (1998), Science publication indicators for India: Questions of interpretation. In: *Proceedings of the 5th International Conference on Science Indicators*, pp. 25-26. Cambridge, CWTS-SPRU.

BONITZ, M. (1999), The Matthew index, concentration patterns and Matthew core journals, *Scientometrics,* 44 (3) : 361–378.

BOOKSTEIN, A. (1990), Infometric distributions, Part II: Resilience to ambiguity, *Journal of the American Society for Information Science*, 41 (5) : 376–386.

EGGHE, L. (1991), The exact place of Zipf's and Pareto's law amongst the classical informetric law, *Scientometrics*, 20 (1) : 93–106.

GARFIELD, E. (1955), Citation indexes for science, *Science*, 122 (3159) : 108–111.

GIBBONS, M. (1999), Science's new social contract with society, *Nature*, 402 : C81-84.

HAITUN, S. D. (1982), Stationary scientometric distributions, Part III. The role of the Zipf distribution, *Scientometrics*, 4 (3) : 181–194.

KATZ, S. (1999), The self-similar science system, *Research Policy*, 28 : 501–517.

KOSTOFF, R. N. (1997), Citation analysis cross-field normalization, a new paradigm, *Scientometrics*, 39 (3) : 225–230.

KOSTOFF, R. N. (1998), The under-reporting of research impact, *The Scientist*, 12 (18) : 9.

LEYDESDORF, L., S. COZZENS (1993), The delineation of specialties in terms of journals using the dynamic journal set of the SCI, *Scientometrics*, 26 (1) : 135–156.

MOED, H. F. (1996), Differences in the construction of SCI based bibliometric indicators among various producers: A first overview, *Scientometrics*, 35 (2) : 177–191.

MORAVCSIK, M. J. (1988), The coverage of science in the Third World: The Philadelphia program. In: EGGHE L., ROUSSEAU R. (Eds), *Informetrics*, 87/88 : 147–155. Amsterdam: Elsevier.

MURUGESAN, P., M. J MORAVCSIK (1978), Variation of the nature of citation measures with journals and scientific disciplines, *Journal of the American Society for Information Science*, 28 : 141–155.

NARIN, F. (1976), *Evaluative Bibliometrics: the Use of Publication and Citation Analysis in the Evaluation of Scientific Activity*, NSF Contract C-627, 456 p.

PRICE, D. J. DE SOLLA (1976), A general theory of bibliometric and other cumulative advantage processes, *Journal of the American Society for Information Science*, 27 : 292–306.

SCHUBERT, A., T. BRAUN (1986), Relative indicators and relational charts for comparative assessment of publication output and citation impact, *Scientometrics*, 9 (5-6) : 281–291.

SCHUBERT, A., T. BRAUN (1996), Cross-field normalization of scientometric indicators, *Scientometrics*, 36 (3) : 311–324.

SIGOGNEAU, A. (2000), Analysis of document types published in journals related to physics, *Scientometrics,* 47 (3) : 589–604.

SIVERTSEN, G. (1992), Should a new bibliometric database for international comparisons be more restricted in journal coverage? In: *Science and Technology in a policy context, Selected Proceedings of the Joint EC-Leiden Conference on Science & Technology Indicators,* DSWO Press, 1992, pp. 35–50.

VAN LEEUWEN, T.N., H. F. MOED, R. J. W. TIJSSEN, M. S. VISSER, A. F. J. VAN RAAN (2001), Language biases in the coverage of science citation index and its consequences for international comparison of national research performance, *Scientometrics*, 51 (1) : 335–346.

VAN RAAN, A. F. J. (2001), Competition among scientists for publication status: toward a model of scientific publication and citation distribution, *Scientometrics*, 51 (1) : 347–357.

WORMELL, I. (1998), Informetric analysis of the international impact of scientific journals: How international are the international journals? *Journal of Documentation*, 54 (5) : 584–605.

ZITT, M., E. BASSECOULARD (1998), Internationalization of scientific journals: A measurement based on publication and citation scope, *Scientometrics*, 41 (1-2) : 255–271.

ZITT, M., F. PERROT, R. BARRÉ (1998), The transition from "national" to "transnational" and related measures of countries' performance, *Journal of the American Society for Information Science*, 49 (1) : 30–42.

ZITT, M., E. BASSECOULARD (1999), Internationalization of communication: A view on the evolution of scientific journals, *Scientometrics*, 46 (3) : 669–685.